

This is a repository copy of *Synthesising quantitative and qualitative evidence to inform guidelines on complex interventions: clarifying the purposes, designs and outlining some methods..*

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/138443/>

Version: Published Version

Article:

Noyes, Jane, Booth, Andrew, Moore, Graham et al. (3 more authors) (2019) Synthesising quantitative and qualitative evidence to inform guidelines on complex interventions: clarifying the purposes, designs and outlining some methods. *BMJ Global health*. ISSN 2059-7908

<https://doi.org/10.1136/bmjgh-2018-000893>

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial (CC BY-NC) licence. This licence allows you to remix, tweak, and build upon this work non-commercially, and any new works must also acknowledge the authors and be non-commercial. You don't have to license any derivative works on the same terms. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Synthesising quantitative and qualitative evidence to inform guidelines on complex interventions: clarifying the purposes, designs and outlining some methods

Jane Noyes,¹ Andrew Booth,² Graham Moore,³ Kate Flemming,⁴ Özge Tunçalp,⁵ Elham Shakibazadeh⁶

To cite: Noyes J, Booth A, Moore G, *et al.* Synthesising quantitative and qualitative evidence to inform guidelines on complex interventions: clarifying the purposes, designs and outlining some methods. *BMJ Glob Health* 2019;4:e000893. doi:10.1136/bmjgh-2018-000893

Handling editor Soumyadeep Bhaumik

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bmjgh-2018-000893>).

Received 11 April 2018
Revised 1 October 2018
Accepted 6 October 2018



©World Health Organization 2019. Licensee BMJ.

For numbered affiliations see end of article.

Correspondence to

Professor Jane Noyes;
jane.noyes@bangor.ac.uk

ABSTRACT

Guideline developers are increasingly dealing with more difficult decisions concerning whether to recommend complex interventions in complex and highly variable health systems. There is greater recognition that both quantitative and qualitative evidence can be combined in a mixed-method synthesis and that this can be helpful in understanding how complexity impacts on interventions in specific contexts. This paper aims to clarify the different purposes, review designs, questions, synthesis methods and opportunities to combine quantitative and qualitative evidence to explore the complexity of complex interventions and health systems. Three case studies of guidelines developed by WHO, which incorporated quantitative and qualitative evidence, are used to illustrate possible uses of mixed-method reviews and evidence. Additional examples of methods that can be used or may have potential for use in a guideline process are outlined. Consideration is given to the opportunities for potential integration of quantitative and qualitative evidence at different stages of the review and guideline process. Encouragement is given to guideline commissioners and developers and review authors to consider including quantitative and qualitative evidence. Recommendations are made concerning the future development of methods to better address questions in systematic reviews and guidelines that adopt a complexity perspective.

INTRODUCTION

Recognition has grown that while quantitative methods remain vital, they are usually insufficient to address complex health systems related research questions.¹ Quantitative methods rely on an ability to anticipate what must be measured in advance. Introducing change into a complex health system gives rise to emergent reactions, which cannot be fully predicted in advance. Emergent reactions can often only be understood through combining quantitative methods with a more flexible

Summary box

- When combined in a mixed-method synthesis, quantitative and qualitative evidence can potentially contribute to understanding how complex interventions work and for whom, and how the complex health systems into which they are implemented respond and adapt.
- The different purposes and designs for combining quantitative and qualitative evidence in a mixed-method synthesis for a guideline process are described.
- Questions relevant to gaining an understanding of the complexity of complex interventions and the wider health systems within which they are implemented that can be addressed by mixed-method syntheses are presented.
- The practical methodological guidance in this paper is intended to help guideline producers and review authors commission and conduct mixed-method syntheses where appropriate.
- If more mixed-method syntheses are conducted, guideline developers will have greater opportunities to access this evidence to inform decision-making.

qualitative lens.² Adopting a more pluralist position enables a diverse range of research options to the researcher depending on the research question being investigated.^{3–5} As a consequence, where a research study sits within the multitude of methods available is driven by the question being asked, rather than any particular methodological or philosophical stance.⁶

Publication of guidance on designing complex intervention process evaluations and other works advocating mixed-methods approaches to intervention research have stimulated better quality evidence for synthesis.^{1 7–13} Methods for synthesising

Box 1 Defining mixed-method research and reviews

Pluye and Hong⁵² define mixed-methods research as “a research approach in which a researcher integrates (a) qualitative and quantitative research questions, (b) qualitative research methods* and quantitative research designs, (c) techniques for collecting and analyzing qualitative and quantitative evidence, and (d) qualitative findings and quantitative results”. A mixed-method synthesis can integrate quantitative, qualitative and mixed-method evidence or data from primary studies.† Mixed-method primary studies are usually disaggregated into quantitative and qualitative evidence and data for the purposes of synthesis. Thomas and Harden further define three ways in which reviews are mixed.⁵³

1. The types of studies included and hence the type of findings to be synthesised (ie, qualitative/textual and quantitative/numerical).
2. The types of synthesis method used (eg, statistical meta-analysis and qualitative synthesis).
3. The mode of analysis: theory testing AND theory building.

*A qualitative study is one that uses qualitative methods of data collection and analysis to produce a narrative understanding of the phenomena of interest. Qualitative methods of data collection may include, for example, interviews, focus groups, observations and analysis of documents.

†The Cochrane Qualitative and Implementation Methods group coined the term ‘qualitative evidence synthesis’ to mean that the synthesis could also include qualitative data. For example, qualitative data from case studies, grey literature reports and open-ended questions from surveys. ‘Evidence’ and ‘data’ are used interchangeably in this paper.

qualitative¹⁴ and mixed-method evidence have been developed or are in development. Mixed-method research and review definitions are outlined in [box 1](#).

This paper is one of a series that aims to explore the implications of complexity for systematic reviews and guideline development, commissioned by WHO. This paper is concerned with the methodological implications of including quantitative and qualitative evidence in mixed-method systematic reviews and guideline development for complex interventions. The guidance was developed through a process of bringing together experts in the field, literature searching and consensus building with end users (guideline developers, clinicians and reviewers). We clarify the different purposes, review designs, questions and synthesis methods that may be applicable to combine quantitative and qualitative evidence to explore the complexity of complex interventions and health systems. Three case studies of WHO guidelines that incorporated quantitative and qualitative evidence are used to illustrate possible uses of mixed-method reviews and mechanisms of integration ([table 1](#), online supplementary files 1–3). Additional examples of methods that can be used or may have potential for use in a guideline process are outlined. Opportunities for potential integration of quantitative and qualitative evidence at different stages of the review and guideline process are presented. Specific considerations when using an evidence to decision framework such as the Developing and Evaluating Communication strategies to support Informed Decisions and practice based on Evidence (DECIDE) framework¹⁵ or the new WHO-INTEGRATE

evidence to decision framework¹⁶ at the review design and evidence to decision stage are outlined. See online supplementary file 4 for an example of a health systems DECIDE framework and Rehfuess *et al*¹⁶ for the new WHO-INTEGRATE framework. Encouragement is given to guideline commissioners and developers and review authors to consider including quantitative and qualitative evidence in guidelines of complex interventions that take a complexity perspective and health systems focus.

Taking a complexity perspective

The first paper in this series¹⁷ outlines aspects of complexity associated with complex interventions and health systems that can potentially be explored by different types of evidence, including synthesis of quantitative and qualitative evidence. Petticrew *et al*¹⁷ distinguish between a complex interventions perspective and a complex systems perspective. A complex interventions perspective defines interventions as having “*implicit conceptual boundaries, representing a flexible, but common set of practices, often linked by an explicit or implicit theory about how they work*”. A complex systems perspective differs in that “*complexity arises from the relationships and interactions between a system’s agents (eg, people, or groups that interact with each other and their environment), and its context. A system perspective conceives the intervention as being part of the system, and emphasises changes and interconnections within the system itself*”. Aspects of complexity associated with implementation of complex interventions in health systems that could potentially be addressed with a synthesis of quantitative and qualitative evidence are summarised in [table 2](#). Another paper in the series outlines criteria used in a new evidence to decision framework for making decisions about complex interventions implemented in complex systems, against which the need for quantitative and qualitative evidence can be mapped.¹⁶ A further paper¹⁸ that explores how context is dealt with in guidelines and reviews taking a complexity perspective also recommends using both quantitative and qualitative evidence to better understand context as a source of complexity. Mixed-method syntheses of quantitative and qualitative evidence can also help with understanding of whether there has been theory failure and or implementation failure. The Cochrane Qualitative and Implementation Methods Group provide additional guidance on exploring implementation and theory failure that can be adapted to address aspects of complexity of complex interventions when implemented in health systems.¹⁹

It may not be apparent which aspects of complexity or which elements of the complex intervention or health system can be explored in a guideline process, or whether combining qualitative and quantitative evidence in a mixed-method synthesis will be useful, until the available evidence is scoped and mapped.^{17 20} A more extensive lead in phase is typically required to scope the available evidence, engage with stakeholders and to refine the review parameters and questions that can then be mapped against potential review designs and methods

Table 1 Designs and methods and their use or applicability in guidelines and systematic reviews taking a complexity perspective

Case study examples and references	Complexity-related questions of interest in the guideline	Types of synthesis used in the guideline	Mixed-method review design and integration mechanisms	Observations, concerns and considerations
A. Mixed-method review designs used in WHO guideline development				
Antenatal Care (ANC) guidelines (online supplementary file 1) 14 15 27 45 48 49	What do women in high-income, medium-income and low-income countries want and expect from antenatal care (ANC), based on their own accounts of their beliefs, views, expectations and experiences of pregnancy? What are the evidence-based practices during ANC that improved outcomes and lead to positive pregnancy experience and how should these practices be delivered? Factors that influence the uptake of routine antenatal services by pregnant women Views and experiences of maternity care providers	Qualitative synthesis Framework synthesis Meta-ethnography Quantitative review of trials Qualitative synthesis Framework synthesis Meta-ethnography	<i>Segregated and contingent design and sequential synthesis</i> Quantitative and qualitative reviews undertaken separately (segregated), an initial scoping review of qualitative evidence established women's preferences and outcomes for ANC, which informed design of the quantitative intervention review (contingent) A second qualitative evidence synthesis was undertaken to look at implementation factors (sequential) Integration: quantitative and qualitative findings were brought together in a series of DECIDE frameworks Tools included: Psychological theory SURE framework conceptual framework for implementing policy options Conceptual framework for analysing integration of targeted health interventions into health systems to analyse contextual health system factors	An innovative approach to guideline development No formal cross-study synthesis process and limited testing of theory. The hypothetical nature of meta-ethnography findings may be challenging for guideline panel members to process without additional training See Flemming <i>et al</i> ¹⁴ for considerations when selecting meta-ethnography
Task shifting guidelines (online supplementary file 2) 15 24 27 45 54	What are the effects of lay health worker interventions in primary and community healthcare on maternal and child health and the management of infectious diseases? What factors affect the implementation of lay health worker programmes for maternal and child health?	Quantitative review of trials Qualitative evidence synthesis Framework synthesis	<i>Combination of a segregated design and sequential synthesis</i> Several published quantitative reviews were used (eg, Cochrane review of lay health worker interventions) Additional new qualitative evidence syntheses were commissioned (segregated) Integration: quantitative and qualitative review findings on lay health workers were brought together in several DECIDE frameworks. Tools included adapted SURE Framework and post hoc logic model	An innovative approach to guideline development The post hoc logic model was developed after the guideline was completed
Risk communication guideline (online supplementary file 3) 15 27 45 55		Quantitative review of quantitative evidence (descriptive) Qualitative using framework synthesis	<i>Results based convergent synthesis</i> A knowledge map of studies was produced to identify the method, topic and geographical spread of evidence. Reviews first organised and synthesised evidence by method-specific streams and reported method-specific findings. Then similar findings across method-specific streams were grouped and further developed using all the relevant evidence Integration: where possible, quantitative and qualitative evidence for the same intervention and question was mapped against core DECIDE domains. Tools included framework using public health emergency model and disaster phases Very few trials were identified. Quantitative and qualitative evidence was used to construct a high level view of what appeared to work and what happened when similar broad groups of interventions or strategies were implemented in different contexts	Example of a fully integrated mixed-method synthesis. Without evidence of effect, it was highly challenging to populate a DECIDE framework
B. Mixed-method review designs that can be used in guideline development				
Factors influencing children's optimal fruit and vegetable consumption ^{30 56}	Potential to explore theoretical, intervention and implementation complexity issues New question(s) of interest are developed and tested in a cross-study synthesis	Mixed-methods synthesis Each review typically has three syntheses: Statistical meta-analysis Qualitative thematic synthesis Cross-study synthesis	<i>Parallel-results convergent synthesis design or a segregated or sequential synthesis design</i> Aim is to generate and test theory from diverse body of literature Integration: used integrative matrix based on programme theory	Can be used in a guideline process as it fits with the current model of conducting method specific reviews separately then bringing the review products together
C. Mixed-method review designs with the potential for use in guideline development				

Continued

Table 1 Continued

Case study examples and references	Complexity-related questions of interest in the guideline	Types of synthesis used in the guideline	Mixed-method review design and integration mechanisms	Observations, concerns and considerations
Interventions to promote smoke alarm ownership and function ^{57 58}	Intervention effect and/or intervention implementation related questions within a system	Narrative synthesis (specifically Popay's methodology) ⁵⁴	<i>Can be accommodated within various review designs</i> Four stage approach to integrate quantitative (trials) with qualitative evidence Integration: initial theory and logic model used to integrate evidence of effect with qualitative case summaries. Tools used included tabulation, groupings and clusters, transforming data: constructing a common rubric, vote-counting as a descriptive tool, moderator variables and subgroup analyses, idea webbing/conceptual mapping, creating qualitative case descriptions, visual representation of relationship between study characteristics and results	Few published examples with the exception of Rodgers, who reinterpreted a Cochrane review on the same topic with narrative synthesis methodology. Methodology is complex. Most subsequent examples have only partially operationalised the methodology. An intervention effect review will still be required to feed into the guideline process
Factors affecting childhood immunisation ⁵⁹	What factors explain complexity and causal pathways?	Bayesian synthesis of qualitative and quantitative evidence	<i>Can be accommodated within various review designs</i> Aim is theory-testing by fusing findings from qualitative and quantitative research Produces a set of weighted factors associated with/predicting the phenomenon under review	Not yet used in a guideline context. Complex methodology. Undergoing development and testing for a health context. The end product may not easily 'fit' into an evidence to decision framework and an effect review will still be required
Providing effective and preferred care closer to home: a realist review of intermediate care. ^{3 60 61}	Developing and testing theories of change underpinning complex policy interventions What works for whom in what contexts and how?	Realist synthesis NB. Other theory-informed synthesis methods follow similar processes	<i>Can be accommodated within various review designs</i> Development of a theory from the literature, analysis of quantitative and qualitative evidence against the theory leads to development of context, mechanism and outcome chains that explain how outcomes come about Integration: programme theory and assembling mixed-method evidence to create Context, Mechanism and Outcome (CMO) configurations	May be useful where there are few trials. The hypothetical nature of findings may be challenging for guideline panel members to process without additional training. The end product may not easily 'fit' into an evidence to decision framework and an effect review will still be required
Use of morphine to treat cancer-related pain ⁶²	Any aspect of complexity could potentially be explored How does the context of morphine use affect the established effectiveness of morphine?	Critical interpretive synthesis	<i>Can be accommodated within various review designs</i> Aims to generate theory from large and diverse body of literature Segregated sequential design Integration: integrative grid	There are few examples and the methodology is complex. The hypothetical nature of findings may be challenging for guideline panel members to process without additional training. The end product would need to be designed to feed into an evidence to decision framework and an intervention effect review will still be required
Food sovereignty, food security and health equity ^{63 64}	Examples have examined health system complexity To understand the state of knowledge on relationships between health equity—ie, health inequalities that are socially produced—and food systems, where the concepts of 'food security' and 'food sovereignty' are prominent Focused on eight pathways to health (in)equality through the food system: (1) Multi-Scalar Environmental, Social Context; (2) Occupational Exposures; (3) Environmental Change; (4) Traditional Livelihoods, Cultural Continuity; (5) Intake of Contaminants; (6) Nutrition; (7) Social Determinants of Health; (8) Political, Economic and Regulatory context	Meta-narrative	Aim is to review research on diffusion of innovation to inform healthcare policy Which research (or epistemic) traditions have considered this broad topic area?; How has each tradition conceptualised the topic (for example, including assumptions about the nature of reality, preferred study designs and ways of knowing)?; What theoretical approaches and methods did they use?; What are the main empirical findings?; and What insights can be drawn by combining and comparing findings from different traditions? Integration: analysis leads to production of a set of meta-narratives ('storylines of research')	Not yet used in a guideline context. The originators are calling for meta-narrative reviews to be used in a guideline process. Potential to provide a contextual overview within which to interpret other types of reviews in a guideline process. The meta-narrative review findings may require tailoring to 'fit' into an evidence to decision framework and an intervention effect review will still be required. Few published examples and the methodology is complex

Table 2 Health-system complexity-related questions that a synthesis of quantitative and qualitative evidence could address (derived from Petticrew *et al*¹⁷)

Aspect of complexity of interest	Examples of potential research question(s) that a synthesis of qualitative and quantitative evidence could address	Types of studies or data that could contribute to a review of qualitative and quantitative evidence
What 'is' the system? How can it be described?	What are the main influences on the health problem? How are they created and maintained? How do these influences interconnect? Where might one intervene in the system?	Quantitative: previous systematic reviews of the causes of the problem); epidemiological studies (eg, cohort studies examining risk factors of obesity); network analysis studies showing the nature of social and other systems Qualitative data: theoretical papers; policy documents
Interactions of interventions with context and adaptation	1. For a research question about Implementation: (How and why) does the implementation of this intervention vary across contexts? 2. For an effectiveness review: Do the effects of the intervention appear to be context dependent?	Qualitative: (1) eg, qualitative studies; case studies Quantitative: (2) trials or other effectiveness studies from different contexts; multicentre trials, with stratified reporting of findings; other quantitative studies that provide evidence of moderating effects of context
System adaptivity (how does the system change?)	(How) does the system change when the intervention is introduced? Which aspects of the system are affected? Does this potentiate or dampen its effects?	Quantitative: longitudinal data; possibly historical data; effectiveness studies providing evidence of differential effects across different contexts; system modelling (eg, agent-based modelling) Qualitative: qualitative studies; case studies
Emergent properties	What are the effects (anticipated and unanticipated) which follow from this system change?	Quantitative: prospective quantitative evaluations; retrospective studies (eg, case-control studies, surveys) may also help identify less common effects; dose-response evaluations of impacts at aggregate level in individual studies or across studies included with systematic reviews (see suggested examples) Qualitative: qualitative studies
Positive (reinforcing) and negative (balancing) feedback loops	What explains change in the effectiveness of the intervention over time? Are the effects of an intervention are damped/suppressed by other aspects of the system (eg, contextual influences?)	Quantitative: studies of moderators of effectiveness; long-term longitudinal studies Qualitative: studies of factors that enable or inhibit implementation of interventions
Multiple (health and non-health) outcomes	What changes in processes and outcomes follow the introduction of this system change? At what levels in the system are they experienced?	Quantitative: studies tracking change in the system over time Qualitative: studies exploring effects of the change in individuals, families, communities (including equity considerations and factors that affect engagement and participation in change)

of synthesis.²⁰ At the scoping stage, it is also common to decide on a theoretical perspective²¹ or undertake further work to refine a theoretical perspective.²² This is also the stage to begin articulating the programme theory of the complex intervention that may be further developed to refine an understanding of complexity and show how the intervention is implemented in and impacts on the wider health system.^{17 23 24} In practice, this process can be lengthy, iterative and fluid with multiple revisions to the review scope, often developing and adapting a logic model¹⁷ as the available evidence becomes known and the potential to incorporate different types of review designs and syntheses of quantitative and qualitative evidence becomes better understood.²⁵ Further questions, propositions or hypotheses may emerge as the reviews progress and therefore the protocols generally need to be developed iteratively over time rather than a priori.

Following a scoping exercise and definition of key questions, the next step in the guideline development process is to identify existing or commission new systematic reviews to locate and summarise the best available evidence in relation to each question. For example, case study 2, 'Optimising health worker roles for maternal and newborn health through task shifting', included quantitative reviews that did and did not take an additional complexity perspective, and qualitative evidence syntheses that were able to explain how specific elements of complexity impacted on intervention outcomes within the wider health system. Further understanding of health system complexity was facilitated through the conduct of additional country-level case studies that contributed to an overall understanding of what worked and what happened when lay health worker interventions were implemented. See [table 1](#) online supplementary file 2.

There are a few existing examples, which we draw on in this paper, but integrating quantitative and qualitative evidence in a mixed-method synthesis is relatively uncommon in a guideline process. **Box 2** includes a set of key questions that guideline developers and review authors contemplating combining quantitative and qualitative evidence in mixed-methods design might ask. Subsequent sections provide more information and signposting to further reading to help address these key questions.

COMPLEXITY-RELATED QUESTIONS THAT A SYNTHESIS OF QUANTITATIVE AND QUALITATIVE EVIDENCE CAN POTENTIALLY ADDRESS

Petticrew *et al*¹⁷ define the different aspects of complexity and examples of complexity-related questions that can potentially be explored in guidelines and systematic reviews taking a complexity perspective. Relevant aspects of complexity outlined by Petticrew *et al*¹⁷ are summarised in [table 2](#) below, together with the corresponding questions that could be addressed in a synthesis combining qualitative and quantitative evidence. Importantly, the aspects of complexity and their associated concepts of

Box 2 Key questions that guideline developers and review authors contemplating combining quantitative and qualitative evidence in a mixed-methods design might ask

1. WHY: Why is a mixed-method synthesis being planned? To answer Compound questions requiring both quantitative and qualitative evidence?
Questions requiring mixed-methods studies?
Separate quantitative and qualitative questions?
2. WHAT: What type of evidence is likely to be available?
Separate quantitative and qualitative research studies?
Related quantitative and qualitative research studies?
Mixed-methods studies?
Quantitative unpublished data and/or qualitative unpublished data, eg, narrative survey data?
3. WHEN: At what point will quantitative and qualitative evidence be integrated?
Throughout the review?
Following separate reviews?
At the question point?
At the synthesis point?
At the evidence to recommendations stage?
Or a combination?
4. HOW: How easy is it to disaggregate quantitative and qualitative data from mixed-method studies? How will quantitative and qualitative evidence be integrated? Through a:
Narrative synthesis or summary?
Quantitising approach, eg, frequency analysis?
Qualitising approach, eg, thematic synthesis?
Tabulation?
Logic model?
Conceptual model/framework?
Matrix?
Graphical approach?
Or a combination?
5. WHICH: Which mixed-method designs, methodologies and methods best fit into a guideline process to inform recommendations?

interest have however yet to be translated fully in primary health research or systematic reviews. There are few known examples where selected complexity concepts have been used to analyse or reanalyse a primary intervention study. Most notable is Chandler *et al*²⁶ who specifically set out to identify and translate a set of relevant complexity theory concepts for application in health systems research. Chandler then reanalysed a trial process evaluation using selected complexity theory concepts to better understand the complex causal pathway in the health system that explains some aspects of complexity in [table 2](#).

Rehfeuss *et al*¹⁶ also recommends upfront consideration of the WHO-INTEGRATE evidence to decision criteria when planning a guideline and formulating questions. The criteria reflect WHO norms and values and take account of a complexity perspective. The framework can be used by guideline development groups as a menu to decide which criteria to prioritise, and which study types and synthesis methods can be used to collect evidence for each criterion. Many of the criteria and their related

questions can be addressed using a synthesis of quantitative and qualitative evidence: the balance of benefits and harms, human rights and sociocultural acceptability, health equity, societal implications and feasibility (see table 3). Similar aspects in the DECIDE framework¹⁵ could also be addressed using synthesis of qualitative and quantitative evidence.

Questions as anchors or compasses

Questions can serve as an ‘anchor’ by articulating the specific aspects of complexity to be explored (eg, Is successful implementation of the intervention context dependent?).²⁷ Anchor questions such as “How does intervention x impact on socioeconomic inequalities in health behaviour/outcome x” are the kind of health system question that requires a synthesis of both quantitative and qualitative evidence and hence a mixed-method synthesis. Quantitative evidence can quantify the difference in effect, but does not answer the question of *how*. The ‘how’ question can be partly answered with quantitative and qualitative evidence. For example, quantitative evidence may reveal where socioeconomic status and inequality emerges in the health system (an emergent property) by exploring questions such as “Does patterning emerge during uptake because fewer people from certain groups come into contact with an intervention in the first place?” or “are people from certain backgrounds more likely to drop out, or to maintain effects beyond an intervention differently?” Qualitative evidence may help understand the reasons behind all of these mechanisms. Alternatively, questions can act as ‘compasses’ where a question sets out a starting point from which to explore further and to potentially ask further questions or develop propositions or hypotheses to explore through a complexity perspective (eg, What factors enhance or hinder implementation?).²⁷ Other papers in this series provide further guidance on developing questions for qualitative evidence syntheses and guidance on question formulation.^{14 28}

For anchor and compass questions, additional application of a theory (eg, complexity theory) can help focus evidence synthesis and presentation to explore and explain complexity issues.^{17 21} Development of a review specific logic model(s) can help to further refine an initial understanding of any complexity-related issues of interest associated with a specific intervention, and if appropriate the health system or section of the health system within which to contextualise the review question and analyse data.^{17 23–25} Specific tools are available to help clarify context and complex interventions.^{17 18}

If a complexity perspective, and certain criteria within evidence to decision frameworks, is deemed relevant and desirable by guideline developers, it is only possible to pursue a complexity perspective if the evidence is available. Careful scoping using knowledge maps or scoping reviews will help inform development of questions that are answerable with available evidence.²⁰ If evidence of effect is not available, then a different approach to develop questions leading to a more general narrative understanding of what happened when complex interventions were implemented

in a health system will be required (such as in case study 3—risk communication guideline). This should not mean that the original questions developed for which no evidence was found when scoping the literature were not important. An important function of creating a knowledge map is also to identify gaps to inform a future research agenda.

Table 2 and online supplementary files 1–3 outline examples of questions in the three case studies, which were all ‘COMPASS’ questions for the qualitative evidence syntheses.

TYPES OF INTEGRATION AND SYNTHESIS DESIGNS IN MIXED-METHOD REVIEWS

The shift towards integration of qualitative and quantitative evidence in primary research has, in recent years, begun to be mirrored within research synthesis.^{29–31} The natural extension to undertaking quantitative or qualitative reviews has been the development of methods for integrating qualitative and quantitative evidence within reviews, and within the guideline process using evidence to decision-frameworks. Advocating the integration of quantitative and qualitative evidence assumes a complementarity between research methodologies, and a need for both types of evidence to inform policy and practice. Below, we briefly outline the current designs for integrating qualitative and quantitative evidence within a mixed-method review or synthesis.

One of the early approaches to integrating qualitative and quantitative evidence detailed by Sandelowski *et al*³² advocated three basic review designs: segregated, integrated and contingent designs, which have been further developed by Heyvaert *et al*³³ (box 3).

A recent review of more than 400 systematic reviews³⁴ combining quantitative and qualitative evidence identified two main synthesis designs—convergent and sequential. In a convergent design, qualitative and quantitative evidence is collated and analysed in a parallel or complementary manner, whereas in a sequential synthesis, the collation and analysis of quantitative and qualitative evidence takes place in a sequence with one synthesis informing the other (box 4).⁶ These designs can be seen to build on the work of Sandelowski *et al*,^{32 35} particularly in relation to the transformation of data from qualitative to quantitative (and vice versa) and the sequential synthesis design, with a cyclical approach to reviewing that evokes Sandelowski’s contingent design.

The three case studies (table 1, online supplementary files 1–3) illustrate the diverse combination of review designs and synthesis methods that were considered the most appropriate for specific guidelines.

METHODS FOR CONDUCTING MIXED-METHOD REVIEWS IN THE CONTEXT OF GUIDELINES FOR COMPLEX INTERVENTIONS

In this section, we draw on examples where specific review designs and methods *have been or can be used* to explore selected aspects of complexity in guidelines or systematic reviews. We also identify other review

Table 3 Integrate evidence to decision framework criteria, example questions and types of studies to potentially address these questions (derived from Rehfeuss *et al*¹⁶)

Domains of the WHO-INTEGRATE EtD framework	Examples of potential research question(s) that a synthesis of qualitative and/or quantitative evidence could address	Types of studies that could contribute to a review of qualitative and quantitative evidence
Balance of benefits and harms	To what extent do patients/beneficiaries <i>value</i> different health outcomes?	Qualitative: studies of views and experiences Quantitative: Questionnaire surveys
Human rights and sociocultural acceptability	Is the intervention <i>socioculturally acceptable</i> to patients/beneficiaries as well as to those implementing it? To what extent do patients/beneficiaries <i>value</i> different non-health outcomes? How does the intervention affect an individual's, population group's or organisation's <i>autonomy</i> , that is, their ability to make a competent, informed and voluntary decision?	Qualitative: discourse analysis, qualitative studies (ideally longitudinal to examine changes over time) Quantitative: pro et contra analysis, discrete choice experiments, longitudinal quantitative studies (to examine changes over time), cross-sectional studies Mixed-method studies; case studies
Health equity, equality and non-discrimination	How <i>affordable</i> is the intervention for individuals, households or communities? How <i>accessible</i> —in terms of physical as well as informational access—is the intervention across different population groups?	Qualitative: studies of views and experiences Quantitative: cross-sectional or longitudinal observational studies, discrete choice experiments, health expenditure studies; health system barrier studies, cross-sectional or longitudinal observational studies, discrete choice experiments, ethical analysis, GIS-based studies
Societal implications	What is the <i>social impact</i> of the intervention: are there features of the intervention that increase or reduce stigma and that lead to social consequences? Does the intervention enhance or limit social goals, such as education, social cohesion and the attainment of various human rights beyond health? Does it change social norms at individual or population level? What is the <i>environmental impact</i> of the intervention? Does it contribute to or limit the achievement of goals to protect the environment and efforts to mitigate or adapt to climate change?	Qualitative: studies of views and experiences Quantitative: RCTs, quasi-experimental studies, comparative observational studies, longitudinal implementation studies, case studies, power analyses, environmental impact assessments, modelling studies
Feasibility and health system considerations	Are there any <i>legal factors</i> that impact on implementation of the intervention? How might <i>governance aspects</i> , such as past decisions and strategic considerations, positively or negatively impact the implementation of the intervention? How does the intervention <i>interact with the existing health system</i> ? Is it likely to fit well or not, is it likely to impact on it in positive or negative ways? How does the intervention interact with the need for and usage of the existing <i>health workforce and broader human resources</i> , at national and subnational levels? How does the intervention interact with the need for and usage of the <i>existing health system infrastructure</i> as well as other relevant infrastructure, at national and subnational levels?	Non-research: policy and regulatory frameworks Qualitative: studies of views and experiences Mixed-method: health systems research, situation analysis, case studies Quantitative: cross-sectional studies

GIS, Geographical Information System; RCT, randomised controlled trial.

Box 3 Segregated, integrated and contingent designs^{32 33}**Segregated design**

Conventional separate distinction between quantitative and qualitative approaches based on the assumption they are different entities and should be treated separately; can be distinguished from each other; their findings warrant separate analyses and syntheses. Ultimately, the separate synthesis results can themselves be synthesised.

Integrated design

The methodological differences between qualitative and quantitative studies are minimised as both are viewed as producing findings that can be readily synthesised into one another because they address the same research purposed and questions. Transformation involves either turning qualitative data into quantitative (quantitising) or quantitative findings are turned into qualitative (qualitising) to facilitate their integration.

Contingent design

Takes a cyclical approach to synthesis, with the findings from one synthesis informing the focus of the next synthesis, until all the research objectives have been addressed. Studies are not necessarily grouped and categorised as qualitative or quantitative.

methods that *could potentially be used* to explore aspects of complexity. Of particular note, we could not find any specific examples of systematic methods to synthesise highly diverse research designs as advocated by Petticrew

*et al*¹⁷ and summarised in tables 2 and 3. For example, we could not find examples of methods to synthesise qualitative studies, case studies, quantitative longitudinal data, possibly historical data, effectiveness studies providing evidence of differential effects across different contexts, and system modelling studies (eg, agent-based modelling) to explore system adaptivity.

There are different ways that quantitative and qualitative evidence can be integrated into a review and then into a guideline development process. In practice, some methods enable integration of different types of evidence in a single synthesis, while in other methods, the single systematic review may include a series of stand-alone reviews or syntheses that are then combined in a cross-study synthesis. Table 1 provides an overview of the characteristics of different review designs and methods and guidance on their applicability for a guideline process. Designs and methods that have already been used in WHO guideline development are described in part A of the table. Part B outlines a design and method that can be used in a guideline process, and part C covers those that have the potential to integrate quantitative, qualitative and mixed-method evidence in a single review design (such as meta-narrative reviews and Bayesian syntheses), but their application in a guideline context has yet to be demonstrated.

Box 4 Convergent and sequential synthesis designs³⁴**Convergent synthesis design**

Qualitative and quantitative research is collected and analysed at the same time in a parallel or complementary manner. Integration can occur at three points:

a. Data-based convergent synthesis design

All included studies are analysed using the same methods and results presented together. As only one synthesis method is used, data transformation occurs (qualitised or quantised). Usually addressed one review question.

b. Results-based convergent synthesis design

Qualitative and quantitative data are analysed and presented separately but integrated using a further synthesis method; eg, narratively, tables, matrices or reanalysing evidence. The results of both syntheses are combined in a third synthesis. Usually addresses an overall review question with subquestions.

c. Parallel-results convergent synthesis design

Qualitative and quantitative data are analysed and presented separately with integration occurring in the interpretation of results in the discussion section. Usually addresses two or more complimentary review questions.

Sequential synthesis design

A two-phase approach, data collection and analysis of one type of evidence (eg, qualitative), occurs after and is informed by the collection and analysis of the other type (eg, quantitative). Usually addresses an overall question with subquestions with both syntheses complementing each other.

POINTS OF INTEGRATION WHEN INTEGRATING QUANTITATIVE AND QUALITATIVE EVIDENCE IN GUIDELINE DEVELOPMENT

Depending on the review design (see boxes 3 and 4), integration can potentially take place at a review team and design level, and more commonly at several key points of the review or guideline process. The following sections outline potential points of integration and associated practical considerations when integrating quantitative and qualitative evidence in guideline development.

Review team level

In a guideline process, it is common for syntheses of quantitative and qualitative evidence to be done separately by different teams and then to integrate the evidence. A practical consideration relates to the organisation, composition and expertise of the review teams and ways of working. If the quantitative and qualitative reviews are being conducted separately and then brought together by the same team members, who are equally comfortable operating within both paradigms, then a consistent approach across both paradigms becomes possible. If, however, a team is being split between the quantitative and qualitative reviews, then the strengths of specialisation can be harnessed, for example, in quality assessment or synthesis. Optimally, at least one, if not more, of the team members should be involved in both quantitative and qualitative reviews to offer the possibility of making connexions throughout the review and not simply at re-agreed junctures. This mirrors O'Cathain's conclusion that mixed-methods primary research tends to work only

when there is a principal investigator who values and is able to oversee integration.^{9 10} While the above decisions have been articulated in the context of two types of evidence, variously quantitative and qualitative, they equally apply when considering how to handle studies reporting a mixed-method study design, where data are usually disaggregated into quantitative and qualitative for the purposes of synthesis (see case study 3—risk communication in humanitarian disasters).

Question formulation

Clearly specified key question(s), derived from a scoping or consultation exercise, will make it clear if quantitative and qualitative evidence is required in a guideline development process and which aspects will be addressed by which types of evidence. For the remaining stages of the process, as documented below, a review team faces challenges as to whether to handle each type of evidence separately, regardless of whether sequentially or in parallel, with a view to joining the two products on completion or to attempt integration throughout the review process. In each case, the underlying choice is of efficiencies and potential comparability vs sensitivity to the underlying paradigm.

Searching

Once key questions are clearly defined, the guideline development group typically needs to consider whether to conduct a single sensitive search to address all potential subtopics (lumping) or whether to conduct specific searches for each subtopic (splitting).³⁶ A related consideration is whether to search separately for qualitative, quantitative and mixed-method evidence ‘streams’ or whether to conduct a single search and then identify specific study types at the subsequent sifting stage. These two considerations often mean a trade-off between a single search process involving very large numbers of records or a more protracted search process retrieving smaller numbers of records. Both approaches have advantages and choice may depend on the respective availability of resources for searching and sifting.

Screening and selecting studies

Closely related to decisions around searching are considerations relating to screening and selecting studies for inclusion in a systematic review. An important consideration here is whether the review team will screen records for all review types, regardless of their subsequent involvement (‘altruistic sifting’), or specialise in screening for the study type with which they are most familiar. The risk of missing relevant reports might be minimised by whole team screening for empirical reports in the first instance and then coding them for a specific quantitative, qualitative or mixed-methods report at a subsequent stage.

Assessment of methodological limitations in primary studies

Within a guideline process, review teams may be more limited in their choice of instruments to assess methodological limitations of primary studies as there are

mandatory requirements to use the Cochrane risk of bias tool³⁷ to feed into Grading of Recommendations Assessment, Development and Evaluation (GRADE)³⁸ or to select from a small pool of qualitative appraisal instruments in order to apply GRADE; Confidence in the Evidence from Reviews of Qualitative Research (GRADE-CERQual)³⁹ to assess the overall certainty or confidence in findings. The Cochrane Qualitative and Implementation Methods Group has recently issued guidance on the selection of appraisal instruments and core assessment criteria.⁴⁰ The Mixed-Methods Appraisal Tool, which is currently undergoing further development, offers a single quality assessment instrument for quantitative, qualitative and mixed-methods studies.⁴¹ Other options include using corresponding instruments from within the same ‘stable’, for example, using different Critical Appraisal Skills Programme instruments.⁴² While using instruments developed by the same team or organisation may achieve a degree of epistemological consonance, benefits may come more from consistency of approach and reporting rather than from a shared view of quality. Alternatively, a more paradigm-sensitive approach would involve selecting the best instrument for each respective review while deferring challenges from later heterogeneity of reporting.

Data extraction

The way in which data and evidence are extracted from primary research studies for review will be influenced by the type of integrated synthesis being undertaken and the review purpose. Initially, decisions need to be made regarding the nature and type of data and evidence that are to be extracted from the included studies. Method-specific reporting guidelines^{43 44} provide a good template as to what quantitative and qualitative data it is potentially possible to extract from different types of method-specific study reports, although in practice reporting quality varies. Online supplementary file 5 provides a hypothetical example of the different types of studies from which quantitative and qualitative evidence could potentially be extracted for synthesis.

The decisions around what data or evidence to extract will be guided by how ‘integrated’ the mixed-method review will be. For those reviews where the quantitative and qualitative findings of studies are synthesised separately and integrated at the point of findings (eg, segregated or contingent approaches or sequential synthesis design), separate data extraction approaches will likely be used.

Where integration occurs during the process of the review (eg, integrated approach or convergent synthesis design), an integrated approach to data extraction may be considered, depending on the purpose of the review. This may involve the use of a data extraction framework, the choice of which needs to be congruent with the approach to synthesis chosen for the review.^{40 45} The integrative or theoretical framework may be decided on a priori if a pre-developed theoretical or conceptual

framework is available in the literature.²⁷ The development of a framework may alternatively arise from the reading of the included studies, in relation to the purpose of the review, early in the process. The Cochrane Qualitative and Implementation Methods Group provide further guidance on extraction of qualitative data, including use of software.⁴⁰

Synthesis and integration

Relatively few synthesis methods start off being integrated from the beginning, and these methods have generally been subject to less testing and evaluation particularly in a guideline context (see [table 1](#)). A review design that started off being integrated from the beginning may be suitable for some guideline contexts (such as in case study 3—risk communication in humanitarian disasters—where there was little evidence of effect), but in general if there are sufficient trials then a separate systematic review and meta-analysis will be required for a guideline. Other papers in this series offer guidance on methods for synthesising quantitative⁴⁶ and qualitative evidence¹⁴ in reviews that take a complexity perspective. Further guidance on integrating quantitative and qualitative evidence in a systematic review is provided by the Cochrane Qualitative and Implementation Methods Group.^{19 27 29 40 47}

Types of findings produced by specific methods

It is highly likely (unless there are well-designed process evaluations) that the primary studies may not themselves seek to address the complexity-related questions required for a guideline process. In which case, review authors will need to configure the available evidence and transform the evidence through the synthesis process to produce explanations, propositions and hypotheses (ie, findings) that were not obvious at primary study level. It is important that guideline commissioners, developers and review authors are aware that specific methods are intended to produce a type of finding with a specific purpose (such as developing new theory in the case of meta-ethnography).⁴⁸ Case study 1 (antenatal care guideline) provides an example of how a meta-ethnography was used to develop a new theory as an end product,^{48 49} as well as framework synthesis which produced descriptive and explanatory findings that were more easily incorporated into the guideline process.²⁷ The definitions ([box 5](#)) may be helpful when defining the different types of findings.

Bringing mixed-method evidence together in evidence to decision (EtD) frameworks

A critical element of guideline development is the formulation of recommendations by the Guideline Development Group, and EtD frameworks help to facilitate this process.¹⁶ The EtD framework can also be used as a mechanism to integrate and display quantitative and qualitative evidence and findings mapped against the EtD framework domains with hyperlinks to more detailed evidence summaries from contributing reviews

Box 5 Different levels of findings

Descriptive findings—qualitative evidence-driven translated descriptive themes that do not move beyond the primary studies.

Explanatory findings—may either be at a descriptive or theoretical level. At the descriptive level, qualitative evidence is used to explain phenomena observed in quantitative results, such as why implementation failed in specific circumstances. At the theoretical level, the transformed and interpreted findings that go beyond the primary studies can be used to explain the descriptive findings. The latter description is generally the accepted definition in the wider qualitative community.

Hypothetical or theoretical finding—qualitative evidence-driven transformed themes (or lines of argument) that go beyond the primary studies. Although similar, Thomas and Harden⁵⁶ make a distinction in the purposes between two types of theoretical findings: analytical themes and the product of meta-ethnographies, third-order interpretations.⁴⁸

Analytical themes are a product of interrogating descriptive themes by placing the synthesis within an external theoretical framework (such as the review question and subquestions) and are considered more appropriate when a specific review question is being addressed (eg, in a guideline or to inform policy).⁵⁶

Third-order interpretations come from translating studies into one another while preserving the original context and are more appropriate when a body of literature is being explored in and of itself with broader or emergent review questions.⁴⁸

(see [table 1](#)). It is commonly the EtD framework that enables the findings of the separate quantitative and qualitative reviews to be brought together in a guideline process. Specific challenges when populating the DECIDE evidence to decision framework¹⁵ were noted in case study 3 (risk communication in humanitarian disasters) as there was an absence of intervention effect data and the interventions to communicate public health risks were context specific and varied. These problems would not, however, have been addressed by substitution of the DECIDE framework with the new INTEGRATE¹⁶ evidence to decision framework. A different type of EtD framework needs to be developed for reviews that do not include sufficient evidence of intervention effect.

DISCUSSION

Mixed-method review and synthesis methods are generally the least developed of all systematic review methods. It is acknowledged that methods for combining quantitative and qualitative evidence are generally poorly articulated.^{29 50} There are however some fairly well-established methods for using qualitative evidence to explore aspects of complexity (such as contextual, implementation and outcome complexity), which can be combined with evidence of effect (see sections A and B of [table 1](#)).¹⁴ There are good examples of systematic reviews that use these methods to combine quantitative and qualitative evidence, and examples of guideline recommendations that were informed by evidence from both quantitative

and qualitative reviews (eg, case studies 1–3). With the exception of case study 3 (risk communication), the quantitative and qualitative reviews for these specific guidelines have been conducted separately, and the findings subsequently brought together in an EtD framework to inform recommendations.

Other mixed-method review designs have potential to contribute to understanding of complex interventions and to explore aspects of wider health systems complexity but have not been sufficiently developed and tested for this specific purpose, or used in a guideline process (section C of table 1). Some methods such as meta-narrative reviews also explore different questions to those usually asked in a guideline process. Methods for processing (eg, quality appraisal) and synthesising the highly diverse evidence suggested in tables 2 and 3 that are required to explore specific aspects of health systems complexity (such as system adaptivity) and to populate some sections of the INTEGRATE EtD framework remain underdeveloped or in need of development.

In addition to the required methodological development mentioned above, there is no GRADE approach³⁸ for assessing confidence in findings developed from combined quantitative and qualitative evidence. Another paper in this series outlines how to deal with complexity and grading different types of quantitative evidence,⁵¹ and the GRADE CERQual approach for qualitative findings is described elsewhere,³⁹ but both these approaches are applied to method-specific and not mixed-method findings. An unofficial adaptation of GRADE was used in the risk communication guideline that reported mixed-method findings. Nor is there a reporting guideline for mixed-method reviews,⁴⁷ and for now reports will need to conform to the relevant reporting requirements of the respective method-specific guideline. There is a need to further adapt and test DECIDE,¹⁵ WHO-INTEGRATE¹⁶ and other types of evidence to decision frameworks to accommodate evidence from mixed-method syntheses which do not set out to determine the statistical effects of interventions and in circumstances where there are no trials.

When conducting quantitative and qualitative reviews that will subsequently be combined, there are specific considerations for managing and integrating the different types of evidence throughout the review process. We have summarised different options for combining qualitative and quantitative evidence in mixed-method syntheses that guideline developers and systematic reviewers can choose from, as well as outlining the opportunities to integrate evidence at different stages of the review and guideline development process.

Review commissioners, authors and guideline developers generally have less experience of combining qualitative and evidence in mixed-methods reviews. In particular, there is a relatively small group of reviewers who are skilled at undertaking fully integrated mixed-method reviews. Commissioning additional qualitative and mixed-method reviews creates an additional cost.

Large complex mixed-method reviews generally take more time to complete. Careful consideration needs to be given as to which guidelines would benefit most from additional qualitative and mixed-method syntheses. More training is required to develop capacity and there is a need to develop processes for preparing the guideline panel to consider and use mixed-method evidence in their decision-making.

CONCLUSION

This paper has presented how qualitative and quantitative evidence, combined in mixed-method reviews, can help understand aspects of complex interventions and the systems within which they are implemented. There are further opportunities to use these methods, and to further develop the methods, to look more widely at additional aspects of complexity. There is a range of review designs and synthesis methods to choose from depending on the question being asked or the questions that may emerge during the conduct of the synthesis. Additional methods need to be developed (or existing methods further adapted) in order to synthesise the full range of diverse evidence that is desirable to explore the complexity-related questions when complex interventions are implemented into health systems. We encourage review commissioners and authors, and guideline developers to consider using mixed-methods reviews and synthesis in guidelines and to report on their usefulness in the guideline development process.

Author affiliations

¹School of Social Sciences, Bangor University, Wales, UK

²School of Health and Related Research (SchHARR), University of Sheffield, Sheffield, UK

³School of Social Sciences, Cardiff University, Wales, UK

⁴Department of Health Sciences, The University of York, York, UK

⁵Department of Reproductive Health and Research including UNDP/UNFPA/UNICEF/WHO/World Bank Special Programme of Research, Development and Research Training in Human Reproduction (HRP), World Health Organization, Geneva, Switzerland

⁶Department of Health Education and Promotion, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran

Contributors JN, AB, GM, KF, ÖT and ES drafted the manuscript. All authors contributed to paper development and writing and agreed the final manuscript. Anayda Portela and Susan Norris from WHO managed the series. Helen Smith was series Editor. We thank all those who provided feedback on various iterations.

Funding Funding provided by the World Health Organization Department of Maternal, Newborn, Child and Adolescent Health through grants received from the United States Agency for International Development and the Norwegian Agency for Development Cooperation.

Disclaimer ÖT is a staff member of WHO. The author alone is responsible for the views expressed in this publication and they do not necessarily represent the decisions or policies of WHO.

Competing interests No financial interests declared. JN, AB and ÖT have an intellectual interest in GRADE CERQual; and JN has an intellectual interest in the iCAT_SR tool.

Patient consent Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Open access This is an open access article distributed under the terms of the Creative Commons Attribution-Non commercial IGO License (CC BY-NC 3.0 IGO), which permits use, distribution, and reproduction for non-commercial purposes in any medium, provided the original work is properly cited. In any reproduction of this article there should not be any suggestion that WHO or this article endorse any specific organization or products. The use of the WHO logo is not permitted. This notice should be preserved along with the article's original URL.

REFERENCES

1. Victora CG, Habicht JP, Bryce J. Evidence-based public health: moving beyond randomized trials. *Am J Public Health* 2004;94:400–5.
2. de Savigny D, Borghi J, Windish R. Chapter 3. Systems thinking: applying a systems perspective to design and evaluate health systems interventions. In: de Savigny D, Tagheed A, eds. *Systems thinking for health systems strengthening*. Geneva: World Health Organisation, 2009.
3. Fletcher A, Jamal F, Moore G, et al. Realist complex intervention science: applying realist principles across all phases of the Medical Research Council framework for developing and evaluating complex interventions. *Evaluation* 2016;22:286–303.
4. Raine R, Fitzpatrick R, Barratt H, et al. Challenges, solutions and future directions in the evaluation of service innovations in health care and public health. *Health Services and Delivery Research* 2016;4:1–136.
5. Baum F. Researching public health: behind the qualitative-quantitative methodological debate. *Soc Sci Med* 1995;40:459–68.
6. Hammersley M. Deconstructing the qualitative quantitative divide. In: Brannen J, ed. *Mixing methods: qualitative and quantitative research*. 1st edn. Aldershot, England: Ashgate, 1992: 39–56.
7. Molina-Azorin JF, Fettes MD. Mixed-methods research prevalence studies: field-specific studies on the state of the art of mixed-methods research. *Journal of Mixed-methods Research* 2016;10:123–8.
8. Moore G, Audrey S. *Process evaluation of complex interventions: medical research council guidance*. MRC Population Health Science Research Network, 2014.
9. O'Cathain A, Hoddinott P, Lewin S, et al. Maximising the impact of qualitative research in feasibility studies for randomised controlled trials: guidance for researchers. *Pilot Feasibility Stud* 2015;1:32.
10. O'Cathain A, Thomas KJ, Drabble SJ, et al. Maximising the value of combining qualitative research and randomised controlled trials in health research: the QUALitative Research in Trials (QUART) study—a mixed methods study. *Health Technol Assess* 2014;18:38.
11. Creswell J, PlanoClark V. *Designing and conducting mixed-methods research*. 2nd edn. USA Sage: Thousand Oaks CA, 2011.
12. Fletcher A, Jamal F, Moore G, et al. Realist complex intervention science: applying realist principles across all phases of the medical research council framework for developing and evaluating complex interventions. *Evaluation* 2016;22:286–303.
13. Creswell JW, Fettes MD, Plano Clark VL. Mixed methods intervention trials. In: Andrew S, Halcomb EJ, eds. *Mixed methods research for nursing and the health sciences*. Oxford: Wiley-Blackwell, 2009: 161–80.
14. Flemming K, Booth A, Garside R, et al. Qualitative evidence synthesis for complex interventions and guideline development: clarification of the purpose, designs and relevant methods. *BMJ Glob Health* 2019;0:e000882. doi:10.1136/bmjgh-2018-000882.
15. DECIDE Evidence to Decision Framework, 2015. Available from: <http://www.decide-collaboration.eu/evidence-decision-etd-framework>
16. Rehfuess EA, Stratil JM, Scheel IB, et al. The WHO integrate evidence to decision framework version 1.0: integrating who norms and values and a complexity perspective. *BMJ Glob Health* 2019;0:e000844. doi:10.1136/bmjgh-2018-000844.
17. Petticrew M, Knai C, Thomas J, et al. Implications of a complexity perspective for systematic reviews and guideline development in health decision making. *BMJ Glob Health* 2019;0:e000899. doi:10.1136/bmjgh-2018-000899.
18. Booth A, Moore G, Flemming K, et al. Taking account of context in systematic reviews and guidelines considering a complexity perspective. *BMJ Glob Health* 2019;0:e000840. doi:10.1136/bmjgh-2018-000840.
19. Cargo M, Harris J, Pantoja T, et al. Cochrane qualitative and implementation methods group guidance series-paper 4: methods for assessing evidence on intervention implementation. *J Clin Epidemiol* 2018;97:59–69.
20. Kelly MP, Noyes J, Kane RL, et al. AHRQ series on complex intervention systematic reviews-paper 2: defining complexity, formulating scope, and questions. *J Clin Epidemiol* 2017;90:11–18.
21. Noyes J, Hendry M, Booth A, et al. Current use was established and Cochrane guidance on selection of social theories for systematic reviews of complex interventions was developed. *J Clin Epidemiol* 2016;75:78–92.
22. Carroll C, Booth A, Leaviss J, et al. "Best fit" framework synthesis: refining the method. *BMC Med Res Methodol* 2013;13:37.
23. Anderson LM, Petticrew M, Rehfuess E, et al. Using logic models to capture complexity in systematic reviews. *Res Synth Methods* 2011;2:33–42.
24. Kneale D, Thomas J, Harris K. Developing and optimising the use of logic models in systematic reviews: exploring practice and good practice in the use of programme theory in reviews. *PLoS One* 2015;10:e0142187.
25. Turley R, Saith R, Bhan N, et al. Slum upgrading review: methodological challenges that arise in systematic reviews of complex interventions. *J Public Health* 2013;35:171–5.
26. Chandler J, Rycroft-Malone J, Hawkes E, et al. Application of simplified complexity theory concepts for healthcare social systems to explain the implementation of evidence into practice. *J Adv Nurs* 2016;72:461–80.
27. Dixon-Woods M. Using framework-based synthesis for conducting reviews of qualitative studies. *BMC Med* 2011;9:39.
28. Booth A, Noyes J, Flemming K, et al. Formulating questions to explore complex interventions within qualitative evidence synthesis. *BMJ Glob Health* 2019;0:e001107. doi:10.1136/bmjgh-2018-001107.
29. Harden A, Thomas J, Cargo M, et al. Cochrane qualitative and implementation methods group guidance series-paper 5: methods for integrating qualitative and implementation evidence within intervention effectiveness reviews. *J Clin Epidemiol* 2018;97:70–8.
30. Thomas J, Harden A, Oakley A. Integrating Qualitative Research with trials in systematic reviews: an example review from public health shows how integration is possible and some potential benefits. *BMJ* 2004;328:1010–2.
31. Petticrew M, Rehfuess E, Noyes J, et al. Synthesizing evidence on complex interventions: how meta-analytical, qualitative, and mixed-method approaches can contribute. *J Clin Epidemiol* 2013;66:1230–43.
32. Sandelowski M, Voils CI, Barroso J. Defining and designing mixed research synthesis studies. *Res Sch* 2006;13:29.
33. Heyvaert M, Hannes K, Onghena P. *Using mixed-methods research synthesis for literature reviews*. New York: Sage Publications, 2016.
34. Hong QN, Pluye P, Bujold M, et al. Convergent and sequential synthesis designs: implications for conducting and reporting systematic reviews of qualitative and quantitative evidence. *Syst Rev* 2017;6:61.
35. Sandelowski M, Voils CI, Leeman J, et al. Mapping the mixed methods-mixed research synthesis terrain. *J Mix Methods Res* 2012;6:317–31.
36. Weir MC, Grimshaw JM, Mayhew A, et al. Decisions about lumping vs. splitting of the scope of systematic reviews of complex interventions are not well justified: a case study in systematic reviews of health care professional reminders. *J Clin Epidemiol* 2012;65:756–63.
37. Higgins JPT, Altman DG, Gotzsche PC, et al. The Cochrane collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343(oct18 2):d5928.
38. GRADE. Grading of Recommendations Assessment, Development and Evaluation GRADE working group. Available from: <http://www.gradeworkinggroup.org/>
39. Lewin S, Glenton C, Munthe-Kaas H, et al. Using qualitative evidence in decision making for health and social interventions: an approach to assess confidence in findings from qualitative evidence syntheses (GRADE-CERQual). *PLoS Med* 2015;12:e1001895.
40. Noyes J, Booth A, Flemming K, et al. Cochrane qualitative and implementation methods group guidance series-paper 3: methods for assessing methodological limitations, data extraction and synthesis, and confidence in synthesized qualitative findings. *J Clin Epidemiol* 2018;97:49–58.
41. Pluye P, Robert E, Cargo M, 2011. Mixed-methods appraisal tool (MMAT). Available from: <http://mixedmethodsappraisaltoolpublic.pbworks.com>
42. CASP. Critical Appraisal Skills Programme (CASP). Qualitative checklist, 2013. Available from: <http://www.casp-uk.net/casp-tools-checklists>
43. Consolidated Standards of Reporting Trials (CONSORT) Statement, 2010. Available from: <http://www.consort-statement.org/>

44. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care* 2007;19:349–57.
45. Higgins JPT, Green S. *Cochrane handbook for systematic reviews of interventions*. Version 5.1.0 updated March 2011, 2011.
46. Higgins JPT, López-López JA, Becker BJ, et al. Synthesising quantitative evidence in systematic reviews of complex health interventions. *BMJ Glob Health* 2019;0:e000858. doi: 10.1136/bmjgh-2018-000858.
47. Flemming K, Booth A, Hannes K, et al. Cochrane qualitative and implementation methods group guidance series-paper 6: reporting guidelines for qualitative, implementation, and process evaluation evidence syntheses. *J Clin Epidemiol* 2018;97:79–85.
48. Noblit GW, Hare D. *Meta-ethnography: synthesizing qualitative studies*. Newbury Park: Sage Publications, 1988.
49. Atun R, de Jongh T, Secci F, et al. Integration of targeted health interventions into health systems: a conceptual framework for analysis. *Health Policy Plan* 2010;25:104–11.
50. Tricco AC, Antony J, Soobiah C, et al. Knowledge synthesis methods for integrating qualitative and quantitative data: a scoping review reveals poor operationalization of the methodological steps. *J Clin Epidemiol* 2016;73:29–35.
51. Montgomery P, Movsisyan A, Sean Grant S, et al. Considerations of complexity in rating certainty of evidence in systematic reviews: a primer on using the grade approach in global health. *BMJ Glob Health* 2019;0:e000848. doi:10.1136/bmjgh-2018-000848.
52. Pluye P, Hong QN. Combining the power of stories and the power of numbers: mixed methods research and mixed studies reviews. *Annu Rev Public Health* 2014;35:29–45.
53. Thomas J, Harden A. Mixed methods and systematic reviews: examples and emerging issues. In: Tashakkori A, Teddlie C, eds. *SAGE handbook of mixed methods in social & behavioral research*. 2nd edn. London: Sage Publications, 2010: 749–74.
54. World Health Organization. SURE Guides for preparing and using evidence-based policy briefs:5. Identifying and addressing barriers to implementing policy options. Available from: <http://www.who.int/evidence/sure/guides/en/>
55. WHO. Framework of phases of disaster response in public health emergencies: prevention; readiness/preparedness; response; recovery, and evaluation. Available from: <http://www.who.int/risk-communication/guidance/process/Q12-Final-Report.pdf?ua=1>
56. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Med Res Methodol* 2008;8:1–10.
57. Popay J, Roberts H, Sowden A, 2007. Guidance on the conduct of narrative synthesis in systematic reviews. Available from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.178.3100&rep=rep1&type=pdf>
58. Rodgers M, Sowden A, Petticrew M. Testing methodological guidance on the conduct of narrative synthesis in systematic reviews: effectiveness of interventions to promote smoke alarm ownership and function. *Evaluation* 2009;15:49–73.
59. Roberts KA, Dixon-Woods M, Fitzpatrick R, et al. Factors affecting uptake of childhood immunisation: a Bayesian synthesis of qualitative and quantitative evidence. *Lancet* 2002;360:1596–9.
60. Pearson M, Hunt H, Cooper C, et al. Providing effective and preferred care closer to home: a realist review of intermediate care. *Health Soc Care Community* 2015;23:577–93.
61. Coryn CLS, Noakes LA, Westine CD, et al. A Systematic review of theory-driven evaluation practice from 1990 to 2009. *Am J Eval* 2011;32:199–226.
62. Flemming K. Synthesis of quantitative and qualitative research: an example using Critical Interpretive Synthesis. *J Adv Nurs* 2010;66:201–17.
63. Greenhalgh T, Robert G, Macfarlane F, et al. Storylines of research in diffusion of innovation: a meta-narrative approach to systematic review. *Soc Sci Med* 2005;61:417–30.
64. Weiler AM, Hergesheimer C, Brisbois B, et al. Food sovereignty, food security and health equity: a meta-narrative mapping exercise. *Health Policy Plan* 2015;30:1078–92.